

Measuring with correlated detectors: violation of Heisenberg noise-disturbance principle and other surprises

Antonio Di Lorenzo

*Instituto de Física, Universidade Federal de Uberlândia,
38400-902 Uberlândia, Minas Gerais, Brazil*

Heisenberg formulated an uncertainty principle stating that there is a tradeoff between noise and disturbance when a measurement of position and a measurement of momentum are performed *sequentially*, and another principle imposing a limitation on the product of the uncertainties in a *joint* measurement of position and momentum. Heisenberg provided semiquantitative examples to support his conjecture, but no proof. Are these formulations of the uncertainty relation sound?

Here, we answer this question by analyzing two nondemolition measurements of position and momentum and describing the quantum state of the detectors. The former literature neglected the possibility of pre-existing correlations between the two detectors. In case these are initially correlated, we prove the Heisenberg noise-disturbance principle not to hold. Furthermore, Ozawa and Arthurs-Kelly inequalities are also violated. Finally, we show that, by preceding a momentum measurement with a position measurement, the noise cancels out if the two probes are in an appropriate EPR state.

PACS numbers: 03.65.Ta, 03.67.-a, 05.40.Ca, 06.20.Dk

Keywords: Uncertainty principle, Noise, Disturbance, Quantum nondemolition measurement

Introduction. The questions that we ask (and answer) in this Letter are of fundamental importance: Does the noise-disturbance principle hold, i.e., is it true that increasing the precision in a measurement of position disturbs in an uncontrollable way the momentum? How do we quantify the disturbance? What happens if we measure position and momentum not sequentially, but jointly?

These questions have been investigated theoretically for more than two score years by a minority of physicists, but recent experiments [1–3] have brought them to the attention of the general public.

Here, we apply Bohr’s prescription of only asking questions that can be answered by experiments, however idealized these may be. To this end, we consider two non-demolition measurements of position and momentum, and describe both the probes and the system quantum-mechanically. Contrary to previous works, we allow the probes to be initially correlated.

The additional term due to the correlations has surprising consequences:

- 1) It makes possible to violate the Heisenberg uncertainty principle, in the noise-disturbance formulation [4].
- 2) It voids a recent noise-disturbance relation derived by Ozawa [5].
- 3) It lowers, from \hbar^2 to $\hbar^2/4$, the limit established by Arthurs and Kelly [6] for the joint uncertainty principle.
- 4) It allows, paradoxically, to counteract the noise introduced in a momentum measurement by preceding it with a position measurement, and vice versa.

The results presented herein are valid for joint or sequential measurements, for any initial preparation of the probes and of the system, and for arbitrary coupling strengths.

Heisenberg [4] formulated the uncertainty principle in

an ambiguous way: either a measurement of position with an uncertainty ϵ_X entails a subsequent disturbance on the momentum $\eta_{P|X}$ such that $\epsilon_X \eta_{P|X} \gtrsim \hbar$, or a joint determination of position and momentum has uncertainties $\epsilon_X \epsilon_P \gtrsim \hbar$. Furthermore, it is unclear whether Heisenberg was referring to the uncertainty ϵ on the subensemble of particles for which a given outcome was found, or to the uncertainty Δ over the whole ensemble. Soon, Kennard and Weyl [7, 8] proved an inequality for the uncertainties of position and momentum, but with a fundamental difference from Heisenberg: the Kennard inequality $\sigma_X \sigma_P \geq \hbar/2$ refers to the uncertainties σ_X, σ_P that one would obtain by measuring X ideally (i.e. with a detector that introduces no noise) on an ensemble of identically prepared particles, and then by measuring P on a distinct ensemble that is prepared in the same way and that has not undergone the X measurement.

The noise-disturbance principle, the joint uncertainty principle, and the Kennard inequality were confused with one another for a long time. A breakthrough came with a paper by Arthurs and Kelly [6], who considered a joint measurement of position and momentum, assuming a von Neumann protocol and probes prepared in a pure Gaussian state. As Arthurs and Kelly considered the system to be initially uncorrelated from the probes, not only they found that Heisenberg principle for joint measurement held, but that the quantum nature of the probes increased the total spread of the outputs, yielding $\Delta_X \Delta_P \geq \sigma_X \sigma_P + \hbar/2 \geq \hbar$. Here, σ refers to the intrinsic spread of the system before the measurement, and Δ to the spread of the pointer variables of the probes after the measurement. We call this relation the Arthurs-Kelly inequality. Although with a two-decade delay, the pioneering result of Arthurs and Kelly stimulated further investigations on the uncertainty principle for joint mea-

measurements of noncommuting variables [9–23].

The distinction between Heisenberg noise-disturbance principle and Kennard inequality became gradually clear [24], and it has given rise to a parallel line of research [5, 25–31].

General theory of quantum measurement. We call measurement any act of inference, even though in a probabilistic sense, about a system through the observation of a second system, the probe, that has interacted with the former. What sets apart the probe from the system is its ability to cause a sensation in a human being, if not directly, through a chain of amplifications and further interactions with the visible electromagnetic field. Here, the effect of this von Neumann chain [32] is accounted for by considering the readout state of the probe to be a mixed state. For clarity, let us suppose that system and probe are initially uncorrelated, so that the initial state is $\hat{\rho}_{pr} \otimes \hat{\rho}_{sys}$. The probability of obtaining an outcome μ from the probe is given by Born's rule

$$P(\mu) = \text{Tr}\{(\hat{\rho}'_{pr}(\mu) \otimes \mathbb{1})U(\hat{\rho}_{pr} \otimes \hat{\rho}_{sys})U^\dagger\}, \quad (1)$$

where U is the time evolution operator, and $\hat{\rho}'_{pr}(\mu)$ are a family of readout density matrices that account for the rest of the von Neumann chain. They must satisfy the normalization $\int d\mu \hat{\rho}'_{pr}(\mu) = 1$, with $d\mu$ a Lebesgue-Stieltjes measure, corresponding to a discrete or continuous distribution of outputs μ . Furthermore, it is sensible to assume that the readout states are classical, i.e. $[\hat{\rho}'_{pr}(\mu_1), \hat{\rho}'_{pr}(\mu_2)] = 0$, $\forall \mu_1, \mu_2$. This hypothesis allows to individuate one, or more, privileged basis, $|J\rangle$, the one that diagonalizes at once all $\hat{\rho}'_{pr}(\mu)$. The existence of such a basis can be justified by the decoherence approach [33, 34]. The $\hat{\rho}'_{pr}(\mu)$ are labeled by the average of the operator $\hat{J} = \int dJ J |J\rangle\langle J|$ and they have a spread $\delta'(\mu)$ in J , i.e.,

$$\mu = \text{Tr}_{pr}\{\hat{J}\hat{\rho}'_{pr}(\mu)\}, \quad (2)$$

$$\delta'^2(\mu) = \text{Tr}_{pr}\{\hat{J}^2\hat{\rho}'_{pr}(\mu)\} - \left(\text{Tr}_{pr}\{\hat{J}\hat{\rho}'_{pr}(\mu)\}\right)^2. \quad (3)$$

The spread $\delta'(\mu)$ is but the resolution of the probe, and in principle it can be different for different outputs.

For instance, we could choose

$$\hat{\rho}'_{pr}(\mu_n) = \int_{\mu_n - \delta'_n/2}^{\mu_n + \delta'_n/2} dJ |J\rangle\langle J| \quad (4)$$

and let the readout μ take discrete values μ_n spaced by $(\delta'_n + \delta'_{n+1})/2$ from one another (then $d\mu$ is a distribution formed by a sum of Dirac deltas), or we could choose

$$\hat{\rho}'_{pr}(\mu) = \int_{-\infty}^{+\infty} dJ \frac{\exp[-(J - \mu)^2/2\delta'^2]}{\sqrt{2\pi}\delta'} |J\rangle\langle J|, \quad (5)$$

and let $\mu \in \mathbb{R}$. In general, we can write

$$\hat{\rho}'_{pr}(\mu) = \int dJ p(\mu|J) |J\rangle\langle J|, \quad (6)$$

with $p(\mu|J)$ a probability distribution for μ . Assuming a uniform resolution $\delta'(\mu) = \delta'$, Eq. (2) is satisfied for $p(\mu|J) = f(\mu - J)$, where $f(\mu)$ is a probability distribution having zero average and spread δ' .

The conditional state of the system for given μ is

$$\hat{\rho}_{sys|\mu} = P(\mu)^{-1} \text{Tr}_{pr}\{(\hat{\rho}'_{pr}(\mu) \otimes \mathbb{1})U(\hat{\rho}_{pr} \otimes \hat{\rho}_{sys})U^\dagger\}. \quad (7)$$

Of course, equivalent formulas arise when treating post-selected weak measurement [35, 36], but with the difference that U is expanded in a perturbation series and the rôles of system and probe are reversed. Let us introduce the preparation basis $|I\rangle$, the one in which $\hat{\rho}_{pr} = \int dI w(I) |I\rangle\langle I|$. We can rewrite the conditional state of the system

$$\hat{\rho}_{sys|\mu} = \frac{\int dJ dI M_{J,I}(\mu) \hat{\rho}_{sys} M_{J,I}^\dagger(\mu)}{\int dJ dI \text{Tr}_{sys}\{E_{J,I}(\mu) \hat{\rho}_{sys}\}}. \quad (8)$$

We defined the generalized operations

$$M_{J,I}(\mu) = \sqrt{p(\mu|J)w(I)} \langle J | \hat{U} | I \rangle, \quad (9)$$

and the generalized effects $E_{J,I}(\mu) = M_{J,I}^\dagger(\mu) M_{J,I}(\mu)$. Both $M_{J,I}(\mu)$ and $E_{J,I}(\mu)$ are operators on the Hilbert space of the system alone. If we do not make the hypothesis of classical readout, Eq. (9) becomes $M_{J,I}(\mu) = \sqrt{p(\mu|J)w(I)} \langle \mu : J | \hat{U} | I \rangle$, where $|\mu : J\rangle$ is the basis that diagonalizes $\hat{\rho}'_{pr}(\mu)$.

When both $\hat{\rho}'_{pr}(\mu)$ and $\hat{\rho}_{pr}$ are pure states, the measurement is positive-operator valued [37–42]. In the general case, we can still define positive operators associated with the measurement

$$E(\mu) = \int dJ dI E_{J,I}(\mu). \quad (10)$$

The operators $E(\mu)$ can be inferred from the observed probabilities $P(\mu)$ by changing the preparation $\hat{\rho}_{sys}$. Recovering $E_{J,I}(\mu)$, however, is a nontrivial task. If $E_{J,I}(\mu)$ are known, the operations are determined modulo a unitary operator: $M_{J,I}(\mu) = V_{J,I}(\mu) E_{J,I}^{1/2}(\mu)$, where $E_{J,I}^{1/2}$ is univocally defined as a positive operator. The additional operations $V_{J,I}(\mu)$ represent an unwanted and avoidable feedback on the system. Ideally, $V_{J,I}(\mu) = 1$. Only by making subsequent measurements on the system, is it possible to check whether the measurement process is introducing this unwanted feedback.

Definitions. We specialize the previous general result to a measurement of position and momentum, either sequential or joint. The probe now consists in two probes, one measuring X the other K . We label each probe with the letter corresponding to the variable that it measures, and let $A \in \{X, K\}$ the generic label. Each probe interacts with the system through a potential \hat{H}_A that commutes with the variable \hat{A} . Furthermore, the interaction is assumed instantaneous, for simplicity. Instead of dropping Planck's constant and declaring that we are working

in units of $\hbar = 1$, we introduce the wave number $K = P/\hbar$ and factor out \hbar in the interaction

$$H_{int} = -\hbar \left[\lambda_X \delta(t + \varepsilon) \hat{\Phi}_X \hat{X} + \lambda_K \delta(t - \varepsilon) \hat{\Phi}_K \hat{K} \right]. \quad (11)$$

Here, $\hat{\Phi}_A$ is the variable of the probe A conjugated to \hat{J}_A , satisfying hence $[\hat{\Phi}_A, \hat{J}_A] = i$. For $\varepsilon \rightarrow 0^-$ K is measured first, then X , while for $\varepsilon \rightarrow 0^+$ the order is exchanged, and finally for $\varepsilon = 0$ the measurements are joint, *à la* Arthurs and Kelly [6]. We shall absorb the coupling constants through the canonical scaling $\Phi_A \rightarrow \lambda_A \Phi_A$, $J_A \rightarrow J_A/\lambda_A$. Thus, J_A has the same dimensions as A , and Φ_A has the inverse dimensions.

We define the initial intrinsic variances of the system $\sigma_A^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$, the initial variances of the probes $\delta_A^2 = \langle \hat{J}_A^2 \rangle - \langle \hat{J}_A \rangle^2$ and $\tilde{\delta}_A^2 = \langle \hat{\Phi}_A^2 \rangle - \langle \hat{\Phi}_A \rangle^2$, their biases $\langle J_A \rangle = \text{Tr}\{\hat{J}_A \hat{\rho}_{pr}\}$, $\langle \Phi_A \rangle = \text{Tr}\{\hat{\Phi}_A \hat{\rho}_{pr}\}$, and their cross-covariances

$$\xi = \langle \hat{\Phi}_X \hat{J}_K \rangle - \langle \hat{\Phi}_X \rangle \langle \hat{J}_K \rangle, \kappa = \langle \hat{\Phi}_K \hat{J}_X \rangle - \langle \hat{\Phi}_K \rangle \langle \hat{J}_X \rangle. \quad (12)$$

Notice that the operators in the covariances commute, so there is no ambiguity. Recall that Kennard inequality establishes that $\sigma_X \sigma_K \geq 1/2$ and $\delta_A \tilde{\delta}_A \geq 1/2$, and that the covariances obey the Cauchy-Schwarz inequalities $|\xi| \leq \tilde{\delta}_X \delta_K$ and $|\kappa| \leq \tilde{\delta}_K \delta_X$.

For later convenience, we introduce the column vectors (indicated by horizontal lists within square brackets) $s = [x, k]$, $\phi = [\phi_K, \phi_X]$, $j = [j_K, j_X]$. The components of the vectors are labeled by the index A , $s_K = x$ and $s_X = k$. We shall make use of the matrices

$$\alpha_+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \alpha_- = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \quad \alpha_0 = \frac{\alpha_+ + \alpha_-}{2}. \quad (13)$$

The Moyal quasi-characteristic function $M(x, k)$ [43] is defined as the Fourier transform of the Wigner quasiprobability $W(K, X)$ [44]. Let $M_{sys}(s)$ the initial Moyal function of the system, and $M_{pr}(\phi, j)$ that of the probes. If the probes are initially uncorrelated, then $M_{pr}(\phi, j) = M_X(\phi_X, j_X) M_K(\phi_K, j_K)$. However, we shall consider the general case.

Noise and disturbance: operational definitions. We wish to discuss the uncertainty relation in the Heisenberg formulation $\epsilon_A \eta_{A'|A} \geq r$, where r is a positive real number ($r \simeq 1$ in the argumentation of Heisenberg), ϵ_A is the noise in the variable that is measured first, and $\eta_{A'|A}$ is the disturbance introduced in the second measurement by the first one. Thus, we need to agree on the definition of ϵ_A and $\eta_{A'|A}$. For definiteness, say that we measure first X , then K .

We shall consider the noise ϵ_X that is introduced by the measuring apparatus, not that intrinsic to the system. We can determine ϵ_X by calibration: the system is prepared in ρ_{sys} and it is measured by a reference probe that introduces no bias and a known noise $\epsilon_X^{(0)}$. By measuring the average, we know the intrinsic average $\langle X \rangle_0$

of the system, and by estimating total variance of the outcome $(\Delta_X^{(0)})^2$, we infer the intrinsic uncertainty of the system $\sigma_X^2 = (\Delta_X^{(0)})^2 - (\epsilon_X^{(0)})^2$. Then we repeat the measurement with the probe that we want to characterize on an identically prepared system. The readout of the probe gives an average $\langle X \rangle$ and a variance Δ_X^2 . We define the statistical noise $\epsilon_X^2 = \Delta_X^2 - \sigma_X^2$ and the systematic error $D_X = \langle X \rangle - \langle X \rangle_0$. The total error $\mathcal{E}_X = (D_X^2 + \epsilon_X^2)^{1/2}$ is the sum in quadrature of the two errors. It may be controversial whether the systematic error would be included by Heisenberg in his formulation, or whether the noise to consider should be only the statistical one. In any case, it is always possible to make D_X zero while leaving ϵ_X unchanged. For a nondemolition measurement, this operational definition coincides with that given by Ozawa.

Notice that in statistics, the statistical noise characterizes the precision, and the systematic error the accuracy. Furthermore ϵ_X represents the uncertainty in each individual measurement, that usually is the predictive uncertainty [17], while Δ_X the ensemble uncertainty [24].

We define the disturbance in an analogous way as we defined the noise. First, we prepare an ensemble of particles in the state ρ_{sys} , skip the X measurement ($\lambda_X = 0$), and measure K . We find an average $\langle K \rangle$ with a statistical spread Δ_K . Then we repeat the procedure, but precede the K -measurement with an X -measurement. We find an average $\langle K \rangle_{|X}$ with a statistical spread $\Delta_{K|X}$. Here, we shall not condition the spread on a specific outcome X_0 of the first measurement, even though we could. We define the statistical disturbance $\eta_{K|X}^2 = \Delta_{K|X}^2 - \Delta_K^2$. Notice that, a priori, η^2 may be negative. A purely imaginary value for η corresponds to the case where the “disturbance” actually decreases the uncertainty, so that it is not a disturbance at all. We also define the systematic disturbance $D_{K|X} = \langle K \rangle_{|X} - \langle K \rangle$. Again, it may be controversial whether this second term represents a disturbance as meant by Heisenberg. Nonetheless, $D_{K|X}$ can be made zero without changing $\eta_{K|X}$. The total disturbance is the sum, in quadrature, of the statistical and systematic disturbances $\mathcal{D}_{K|X} = (\eta_{K|X}^2 + D_{K|X}^2)^{1/2}$. Even for nondemolition measurements, this definition differs from the one given by Ozawa, since the latter does not describe the second measurement process, and thus postulates implicitly that the second probe is making an ideal measurement and is uncorrelated to the first one.

It is important to remark that in principle both the noise and the disturbance could depend on the state of the system. This dependence has been pointed out as undesirable [45], but we see no problem in this: depending how a system is prepared, the effect of a probe on it could be more or less pronounced. Nonetheless, in the following we shall prove that, for the nondemolition measurements considered here, noise and disturbance are properties of the probes alone.

Results. We shall use the relation between the final

Moyal function for system and probes in terms of the initial ones [46, 47]

$$M(s; \phi, j) = M_{sys}(s + \phi) M_{pr}(\phi, j + 2\alpha_0 s + \alpha_\varepsilon \phi). \quad (14)$$

The symbol $\varepsilon \in \{+, -, 0\}$ specifies the order of the measurements, as in Eq. (11). References [46, 47] assumed as readout states $\rho'_{pr}(\mu) = |J = \mu\rangle\langle J = \mu|$. In this case, the characteristic function of the readout probability $Z(\phi) = \int d\mu \exp[iJ \cdot \phi] \Pi(J)$ is obtained by tracing out the system ($s = 0$) and by putting $j = 0$ in Eq. (14). In a realistic measurement with finite resolution, however, $\rho'_{pr}(\mu) = \int dJ p(\mu|J) |J\rangle\langle J|$. We recall that for uniform resolution $p(\mu|J) = f(\mu - J)$. The results of Refs. [46, 47] are then easily generalized thanks to the convolution theorem: the characteristic function, i.e. the Fourier transform of the joint probability of observing $\mu = (\mu_K, \mu_X)$ as the output,

$$Z(\phi) = \int d\mu e^{i\mu \cdot \phi} P(\mu) = \int d\mu dJ e^{i\mu \cdot \phi} p(\mu|J) \Pi(J), \quad (15)$$

is obtained by adding a factor $z(\phi) = \int d\mu \exp[i\mu \cdot \phi] f(\mu)$,

$$Z(\phi) = z(\phi) M_{sys}(\phi) M_{pr}(\phi, \alpha_\varepsilon \phi). \quad (16)$$

For definiteness, we consider $\varepsilon = +$. By taking the second derivatives of $\ln(Z)$, we obtain the statistical variances of the readout

$$\Delta_X^2 = \sigma_X^2 + \delta_X^2 + \delta_X'^2, \quad (17)$$

$$\Delta_{K|X}^2 = \sigma_K^2 + \delta_K^2 + \tilde{\delta}_X^2 + \delta_K'^2 + 2\kappa. \quad (18)$$

The statistical error is $\epsilon_X^2 = \delta_X^2 + \delta_X'^2$, and, since $\Delta_K^2 = \sigma_K^2 + \delta_K^2 + \delta_K'^2$, the statistical disturbance is

$$\eta_{K|X}^2 = \tilde{\delta}_X^2 + 2\kappa. \quad (19)$$

As we noted earlier, we can make the systematic error D_X and disturbance $D_{K|X}$ vanish, by applying two appropriate translations to the initial state of the probe (since we are looking for minimal disturbance and error, it is sensible to assume that the probe X is not biased).

Violation of the noise-disturbance principle. For initially uncorrelated probes $\kappa = 0$, so that $\eta_{K|X}^2 = \tilde{\delta}_X^2$. The Heisenberg noise-disturbance relation becomes then

$$\epsilon_X \eta_{K|X} = \tilde{\delta}_X \sqrt{\delta_X^2 + \delta_X'^2} \geq \frac{1}{2}. \quad (20)$$

Thus, if the probes are not correlated, the noise-disturbance relation for position and momentum holds, in agreement with previous results [25, 27, 45].

However, if the probes are correlated, κ can be negative. The disturbance has the bounds

$$\tilde{\delta}_X(\tilde{\delta}_X - 2\delta_K) \leq \eta_{K|X}^2 \leq \tilde{\delta}_X(\tilde{\delta}_X + 2\delta_K). \quad (21)$$

If $\tilde{\delta}_X \leq 2\delta_K$, the lower limit is negative, so that the disturbance can be arbitrarily small, invalidating Heisenberg's argument. Other possibilities leading to this violation explored in the former literature are the EPR thought-experiment [48], and Ozawa's model [25]. Let us discuss them briefly: in the EPR setup, one particle works as a probe, but the two particles are initially correlated. Thus the hypothesis that system and probes are initially uncorrelated is violated. Furthermore, due to this limitation, the EPR scheme does not work for any initial state of the system. On the other hand, Ozawa violation relies on a single probe whose coordinate y gives information about the momentum p_x of a particle (possibly coinciding with the probe) and whose momentum p_y gives information about x . In this case, one can measure either y or p_y , but not both, hence it is questionable whether Ozawa's model produces a genuine violation of the noise-disturbance principle, as noise and disturbance refer to different setups.

Violation of Ozawa relation. According to Ozawa [5], the following relation holds

$$\epsilon_X \eta_{K|X} + \epsilon_X \sigma_K + \sigma_X \eta_{K|X} \geq \frac{1}{2}. \quad (22)$$

However, we have demonstrated that $\eta_{K|X}$ can be zero. Ozawa's relation then reduces to $\epsilon_X \sigma_K \geq 1/2$. This inequality can be violated, since ϵ_X and σ_K are independent of each other, the first referring to the probe, the second to the system. This does not show that Ozawa's relation is wrong, as it is a mathematical identity, but that the definition of disturbance upon which it relies is incompatible with our operational definition, since it disregards the possibility of correlated probes, and therefore it should be discarded.

Violation of the Arthurs-Kelly limit. Heisenberg also formulated the uncertainty principle in terms of the mutual uncertainty introduced in a joint measurement of momentum and position. Thus, let us consider the case of joint measurements, $\varepsilon = 0$. The variances are

$$\begin{aligned} \Delta_X^2 &= \sigma_X^2 + \delta_X^2 + \delta_X'^2 + \frac{1}{4}\tilde{\delta}_K^2 - \xi \\ &\geq \sigma_X^2 + (\delta_X - \frac{1}{2}\tilde{\delta}_K)^2 + \delta_X'^2, \end{aligned} \quad (23)$$

$$\begin{aligned} \Delta_K^2 &= \sigma_K^2 + \delta_K^2 + \delta_K'^2 + \frac{1}{4}\tilde{\delta}_X^2 + \kappa \\ &\geq \sigma_K^2 + (\delta_K - \frac{1}{2}\tilde{\delta}_X)^2 + \delta_K'^2. \end{aligned} \quad (24)$$

If we interpret Heisenberg's argument as referring to the noise introduced by the measurement, then $\epsilon_X^2 \epsilon_K^2 = (\delta_X^2 + \delta_X'^2 + \frac{1}{4}\tilde{\delta}_K^2 - \xi)(\delta_K^2 + \delta_K'^2 + \frac{1}{4}\tilde{\delta}_X^2 + \kappa)$ has no lower bound. For instance, we could make Φ_K perfectly correlated to J_X , then adjust the couplings so that $2\delta_X = \tilde{\delta}_K$ and let the resolution $\delta_X' \rightarrow 0$. However, if we, as Arthurs and Kelly did [6], interpret Heisenberg's argument as referring to the total uncertainty, that intrinsic to the measured

system plus the noise introduced by the measurement, then $\Delta_X \Delta_K \geq 1/2$. This value is to be contrasted to the one commonly accepted [10] $\Delta_X \Delta_K \geq \sigma_X \sigma_K + 1/2 \geq 1$. The reason for the discrepancy lies once again in the implicit assumption of uncorrelated probes made in the former literature.

Cancelling the noise. Finally, let us go back to the case of sequential measurements, and assume perfect anticorrelations between Φ_X and J_K , so that $\kappa = -\tilde{\delta}_X \delta_K$. For instance, we could prepare the probes in an EPR state. Equation (18) becomes then

$$\Delta_{K|X}^2 = \sigma_K^2 + (\delta_K - \tilde{\delta}_X)^2 + \delta_K'^2. \quad (25)$$

By making $\delta_K = \tilde{\delta}_X$, the contributions of the probes to the variance partially cancel out, leaving only the contribution of the finite resolution, which can be made arbitrarily small, in principle. Remember that we concealed the coupling constants by rescaling the variables. Let us restore them momentarily, so that the equality we want to reach is $\lambda_K^{-1} \delta_K = \lambda_X \tilde{\delta}_X$. This can be realized simply by changing $\lambda_K \lambda_X$, without acting on the initial state of the probes. The same considerations hold when the ordering of the measurements is exchanged.

Conclusions. We have demonstrated how initial correlations in the detectors invalidate many results commonly believed to hold, and how this can be exploited to reduce the measurement noise. The Moyal formalism allowed to unify the treatment of joint and sequential measurements. An interesting question to be investigated is whether classical correlations are sufficient to violate the Heisenberg, Ozawa, and Arthurs-Kelly limits, or whether the detectors should be prepared in an entangled state.

Acknowledgments. I acknowledge discussions with Jacques Distler. This work was performed as part of the Brazilian Instituto Nacional de Ciência e Tecnologia para a Informação Quântica (INCT-IQ).

[1] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, *Nat. Phys.* **8**, 185 (2012).
[2] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, *Phys. Rev. Lett.* **109**, 100404 (2012).
[3] M. M. Weston, M. J. W. Hall, M. S. Pals-son, H. M. Wiseman, and G. J. Pryde, “Experimental test of universal joint measurement uncertainty relations,” (2012), [arXiv:1211.0370](#).
[4] W. Heisenberg, *Zeitschr. f. Phys.* **43**, 172 (1927).
[5] M. Ozawa, *Phys. Rev. A* **67**, 042105 (2003).
[6] E. Arthurs and J. L. Kelly Jr., *Bell System Tech. J.* **44**, 725 (1965).
[7] E. H. Kennard, *Zeitschr. f. Phys.* **44**, 326 (1927).
[8] H. Weyl, *Gruppentheorie und Quantenmechanik* (Hirzel, Leipzig, 1928) [English transl. H. P. Robertson, *The theory of groups and quantum mechanics* (Dover Publications, 1931)].

[9] H. P. Yuen, *Phys. Lett. A* **91**, 101 (1982).
[10] E. Arthurs and M. S. Goodman, *Phys. Rev. Lett.* **60**, 2447 (1988).
[11] H. Martens and W. M. de Muynck, *Found. Phys.* **20**, 357 (1990).
[12] H. Martens and W. M. de Muynck, *Phys. Lett. A* **157**, 441 (1991).
[13] S. Ishikawa, *Rep. Math. Phys.* **29**, 257 (1991).
[14] S. Ishikawa, *Int. J. Theor. Phys.* **30**, 401 (1991).
[15] M. G. Raymer, *Am. J. Phys.* **62**, 986 (1994).
[16] U. Leonhardt and H. Paul, *Prog. Quant. Electron.* **19**, 89 (1995).
[17] D. M. Appleby, *J. Phys. A: Math. Gen.* **31**, 6419 (1998).
[18] A. Trifonov, G. Björk, and J. Söderholm, *Phys. Rev. Lett.* **86**, 4423 (2001).
[19] G. D’Ariano, *Fortschr. Phys.* **51**, 318 (2003).
[20] M. Ozawa, *Interdiscip. Inform. Sci.* **10**, 69 (2004).
[21] M. J. W. Hall, *Phys. Rev. A* **69**, 052113 (2004).
[22] R. F. Werner, *Quantum Info. Comput.* **4**, 546 (2004).
[23] P. Busch, T. Heinonen, and P. Lahti, *Phys. Rep.* **452**, 155 (2007).
[24] L. E. Ballentine, *Rev. Mod. Phys.* **42**, 358 (1970).
[25] M. Ozawa, *Phys. Lett. A* **299**, 1 (2002).
[26] M. Ozawa, *Phys. Lett. A* **318**, 21 (2003).
[27] M. Ozawa, *Ann. Phys.* **311**, 350 (2004).
[28] P. Busch, T. Heinonen, and P. Lahti, *Phys. Lett. A* **320**, 261 (2004).
[29] Y. I. Vorontsov, *Phys.-Usp.* **48**, 999 (2005).
[30] A. Lund and H. Wiseman, *New J. Phys.* **12**, 093011 (2010).
[31] J. Distler and S. Paban, “On Uncertainties in Successive Measurements,” (2012), [arXiv:1211.4169](#).
[32] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932) [*Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1996)].
[33] W. H. Zurek, *Phys. Rev. D* **24**, 1516 (1981).
[34] W. H. Zurek, *Phys. Rev. D* **26**, 1862 (1982).
[35] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. Lett.* **60**, 1351 (1988).
[36] A. Di Lorenzo, *Phys. Rev. A* **85**, 032106 (2012).
[37] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).
[38] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, London, 1976).
[39] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory* (North Holland, Amsterdam, 1982).
[40] K. Kraus, *States, effects, and operations: fundamental notions of quantum theory* (Springer-Verlag, 1983).
[41] G. Ludwig, *Foundations of Quantum Mechanics I* (Springer London, Limited, 1983).
[42] P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics* (Springer, Berlin, 1995) second corrected printing.
[43] J. E. Moyal, *Math. Proc. Cambr. Phil. Soc.* **45**, 99 (1949).
[44] E. Wigner, *Phys. Rev.* **40**, 749 (1932).
[45] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, 1992).
[46] A. Di Lorenzo, *Phys. Rev. A* **83**, 042104 (2011).
[47] A. Di Lorenzo, “Sequential measurement of conjugate variables as an,” (2012), to appear in *Phys. Rev. Lett.*, [arXiv:1205.1238](#).
[48] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).